

\*\*\*\*\*  
W D S E L I F  
\*\*\*\*\* (TW)

Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Msrch\_Pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat May 13 09:51:03 2000; MasPar time 4.55 Seconds  
Tabular output not generated. 327.787 Million cell updates/sec

Title: >US-09-331-631-1  
Description: (186-248) from US09331631.pep (5 of 5)  
Perfect Score: 494  
Sequence: 1 KRDPQREYEDCRRRCRCEOE.....MNPQSGSGRYEGEEQS 63

Scoring table: PAM 150  
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:geneseqp

Statistics: Mean 25.007; Variance 101.250; scale 0.247

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	494	100.0	666	1	Macadamia integrifolia	7.86e-40
2	449	90.9	666	1	Macadamia integrifolia	3.23e-35
3	433	87.7	625	1	Macadamia integrifolia	1.40e-33
4	170	34.4	525	1	Theobroma cacao antlmi	1.94e-07
5	170	34.4	566	1	Sequence encoded by 67	1.94e-07
6	157	31.8	590	1	Gossypium hirsutum ant	3.22e-06
7	141	28.5	637	1	Hordeum vulgare antlmi	9.77e-05
8	115	22.3	593	1	Zea mays antimicrobial	2.17e-02
9	110	22.3	918	1	Human androgen recepto	5.99e-02
10	110	22.3	919	1	Androgen receptor.	5.99e-02
11	110	22.3	919	1	Human androgen recepto	5.99e-02
12	99	20.0	919	1	Human androgen recepto	5.38e-01
13	93	18.8	28	1	Stenocarpus sinuatus a	1.74e+00
14	92	18.6	669	1	Mouse liver cancer-ori	2.11e+00
15	88	17.8	712	1	Rat Y7521 gene product	4.56e+00
16	87	17.6	193	1	Partial human Y7521 ge	5.52e+00
17	85	17.2	371	1	Epitope tagged T8P pro	8.07e+00
18	84	17.0	539	1	Yeast transcriptional	9.75e+00
19	81	16.4	395	1	Mouse SRV-related prot	1.71e+01
20	81	16.4	1382	1	Human metastasis-associ	1.71e+01
21	81	16.4	2289	1	Protein derived from r	1.71e+01
22	80	16.2	33	1	Zea mays antimicrobial	2.06e+01
23	80	16.2	86	1	GST-HD fusion protein	2.06e+01

ID	Score	Query Match	Length	ID	Description	Pred. No.	
24	80	16.2	86	1	W5073	GST-HD fusion protein	2.06e+01
25	80	16.2	436	1	W03662	Human 70K UI snRNP pro	2.06e+01
26	80	16.2	591	1	W48796	Drosophila melanogaster	2.06e+01
27	80	16.2	614	1	R82630	70K autoantigen, part	2.06e+01
28	80	16.2	816	1	R71111	Spinochelliar ataxia	2.06e+01
29	79	16.0	431	1	W52828	Human Nono DNA protein	2.48e+01
30	79	16.0	441	1	R41875	Steroid hormone recept	2.48e+01
31	79	16.0	441	1	R89214	Peroxisome proliferato	2.48e+01
32	79	16.0	971	1	W48896	Candida albicans CaCLA	2.48e+01
33	78	15.8	763	1	W08136	Human cytokine respons	2.99e+01
34	78	15.8	763	1	R49790	Sequence of special AT	2.99e+01
35	78	15.8	763	1	W92415	Human SAT1 protein.	2.99e+01
36	78	15.8	763	1	W26407	Matrix/scarfoid-associ	2.99e+01
37	77	15.6	48	1	W06970	Modified hsp60 gene pr	3.59e+01
38	77	15.6	120	1	W52825	Human TFE3/Nono 5'-RAC	3.59e+01
39	77	15.6	186	1	W26536	Trypanosoma cruzi anti	3.59e+01
40	77	15.6	484	1	P71081	Sequence encoded by Vi	3.59e+01
41	77	15.6	516	1	P61362	Soybean glycinin A3a4	3.59e+01
42	77	15.6	700	1	W52822	Human Nono/TFE3 fusion	3.59e+01
43	76	15.4	198	1	Y04890	Mycobacterium species	4.32e+01
44	76	15.4	198	1	Y04889	Mycobacterium species	4.32e+01
45	76	15.4	205	1	W78191	Human secreted protein	4.32e+01

## ALIGNMENTS

RESULT	ID	Score	Query Match	Length	ID	Description	Pred. No.
1	W62828	standard; Protein; 666 AA.			1	Macadamia integrifolia	7.86e-40
2	W62828	(first entry)			1	Macadamia integrifolia	3.23e-35
3	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
4	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
5	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
6	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
7	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
8	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
9	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
10	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
11	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
12	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
13	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
14	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
15	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
16	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
17	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
18	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
19	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
20	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
21	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
22	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33
23	W62828	antimicrobial protein; infection; control.			1	Macadamia integrifolia	1.40e-33

OS Macadamia integrifolia.  
 FH Key Location/Qualifiers  
 FT Peptide 1..28  
 FT /note="signal peptide"  
 FT Protein 29..666  
 FT /note="mature protein"  
 PN WO9827805-A1.  
 PD 02-JUL-1998.  
 PE 22-DEC-1997; AU0874.  
 PR 20-DEC-1996; AU-004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NT, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DR WPI: 98-377279/32.  
 DR N-PSDB: V42311.  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 useful for controlling microbial infestations of plants or mammals  
 PS Claim 1: Page 39-41: 96pp: English.  
 CC The sequence is that of an antimicrobial protein which can  
 be used to control microbial infestations in plants and mammalian  
 CC animals.  
 CC Sequence 666 AA;

Query Match 90.9%; Score 449; DB 1; Length 666;  
 Best Local Similarity 92.1%; Pred. No. 3.23e-35;  
 Matches 58; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 DB 186 KRDPQREYEDCRRCRCEQEPFOYQRCRCRCEQROHGRGDLINFORGSGRYEEGEE 245  
 QY 186 KRDPQREYEDCRRCRCEQEPFOHQCQLRCRQROHGRGDMNPGRGSGRYEEGEE 245  
 DB 246 KQS 248  
 QY 246 EQS 248

RESULT 3  
 ID W62830 standard; Protein: 625 AA.  
 AC W62830;  
 DT 27-OCT-1998 (first entry)  
 DE Macadamia integrifolia antimicrobial protein.  
 KM antimicrobial protein; infestation; control.  
 OS Macadamia integrifolia.  
 FH Key Location/Qualifiers  
 FT Peptide 1..28  
 FT /note="signal peptide"  
 FT Protein 29..666  
 FT /note="mature protein"  
 PN WO9827805-A1.  
 PD 02-JUL-1998.  
 PE 22-DEC-1997; AU0874.  
 PR 20-DEC-1996; AU-004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NT, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DR WPI: 98-377279/32.  
 DR N-PSDB: V42311.  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 useful for controlling microbial infestations of plants or mammals  
 PS Claim 1: Page 43-45: 96pp: English.  
 CC The sequence is that of an antimicrobial protein which can  
 be used to control microbial infestations in plants and mammalian  
 CC animals.  
 CC Sequence 625 AA;

Query Match 87.7%; Score 433; DB 1; Length 625;  
 Best Local Similarity 88.9%; Pred. No. 1.40e-33;  
 Matches 56; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 DB 145 KRDPQREYEDCRRCRCEQEPFOYQRCRCRCEQROHGRGDLINFORGSGRYEEGEE 204  
 QY 186 KRDPQREYEDCRRCRCEQEPFOHQCQLRCRQROHGRGDMNPGRGSGRYEEGEE 245  
 DB 205 KQS 207

QY 246 EQS 248

RESULT 4  
 ID W62831 standard; Protein: 525 AA.  
 AC W62831;  
 DT 27-OCT-1998 (first entry)  
 DE Theobroma cacao antimicrobial protein.  
 KM antimicrobial protein; infestation; control.  
 OS Theobroma cacao.  
 PN WO9827805-A1.  
 PD 02-JUL-1998.  
 PE 22-DEC-1997; AU0874.  
 PR 20-DEC-1996; AU-004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NT, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DR WPI: 98-377279/32.  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 useful for controlling microbial infestations of plants or mammals  
 PS Claim 1: Page 47-49: 96pp: English.  
 CC The sequence is that of an antimicrobial protein which can  
 be used to control microbial infestations in plants and mammalian  
 CC animals.  
 CC Sequence 525 AA;

Query Match 34.4%; Score 170; DB 1; Length 525;  
 Best Local Similarity 40.7%; Pred. No. 1.94e-07;  
 Matches 24; Conservative 14; Mismatches 18; Indels 3; Gaps 3;  
 DB 82 QROYQCCGRCEQROHGRGDLINFORGSGRYEEGEE 139  
 QY 191 QREYEDCRRCRCEQEPFOHQCQLRCRQROHGRGDMNPGRGSGRYEEGEE 247

RESULT 5  
 ID R20181 standard; Protein: 566 AA.  
 AC R20181;  
 DT 16-APR-1992 (first entry)  
 DE Sequence encoded by 67 kD T. cacao protein cDNA.  
 KM Cocoa; flavour; vicillin; seed storage protein.  
 OS Theobroma cacao.  
 PN WO9119801-A.  
 PD 26-DEC-1991.  
 PE 07-JUN-1991; G00914.  
 PR 11-JUN-1990; GB-013016.  
 PA (MRSC) MARS UK LTD.  
 PI Spencer ME, Hodge R, Deakin EA, Ashton S;  
 DR WPI: 92-024418/03.  
 DR N-PSDB: Q20377.  
 PT Recombinant cocoa proteins - are responsible for flavour in cocoa  
 PT beans and produced in large quantities using yeast and bacterial  
 PT expression vectors  
 PS Claim 4: Fig 2: 59pp: English.  
 CC The inventors claim a 67 kD and 31 kD T. cacao protein, and  
 CC fragments, and encoding DNAs. The 47 kD and 31 kD proteins are  
 CC derived from the 67 kD precursor. T. cacao protein cDNA was  
 CC detected in a cDNA library prepared from immature cocoa beans RNA  
 CC using a probe based on the AA sequence of a CNBr peptide common to  
 CC the 47 kD and 31 kD polypeptides. Homology searches revealed close  
 CC homologues between the 67 kD polypeptide and the vicillins, which are  
 CC seed storage proteins.  
 CC Sequence 566 AA;

Query Match 34.4%; Score 170; DB 1; Length 566;  
 Best Local Similarity 40.7%; Pred. No. 1.94e-07;  
 Matches 24; Conservative 14; Mismatches 18; Indels 3; Gaps 3;  
 DB 82 QROYQCCGRCEQROHGRGDLINFORGSGRYEEGEE 139  
 QY 191 QREYEDCRRCRCEQEPFOHQCQLRCRQROHGRGDMNPGRGSGRYEEGEE 247

RESULT 6

ID W62832 standard; Protein; 590 AA.  
AC W62832;  
DE 27-OCT-1998 (first entry)  
DT Gossypium hirsutum antimicrobial protein.  
KW antimicrobial protein; infestation; control.  
OS Gossypium hirsutum.  
PN W09827805-A1.  
PD 02-JUL-1998.  
PF 22-DEC-1997; AU0874.  
PR (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP.  
DR WPI: 98-377279/32.  
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
PS useful for controlling microbial infestations of plants or mammals  
PS Claim 1; Page 49-51; 96pp; English.  
CC The sequence is that of an antimicrobial protein which can  
CC be used to control microbial infestations in plants and mammalian  
CC animals.  
SQ Sequence 590 AA;  
Query Match 31.8%; Score 157; DB 1; Length 590;  
Best Local Similarity 56.1%; Pred. No. 3, 22e-06;  
Matches 23; Conservative 9; Mismatches 5; Indels 4; Gaps 3;  
Db 82 DPQR-YEECOEQEQRQEQPQCCQCLKRFEEQEQQSQ 121  
QY 188 DPQREYEDCRRCQEQEPRQHQCLRC--R-EQORQHGR 225  
RESULT 7  
ID W62837 standard; Protein; 637 AA.  
AC W62837;  
DE 27-OCT-1998 (first entry)  
DT Hordeum vulgare antimicrobial protein.  
KW antimicrobial protein; infestation; control.  
OS Hordeum vulgare.  
PN W09827805-A1.  
PD 02-JUL-1998.  
PF 22-DEC-1997; AU0874.  
PR (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP.  
DR WPI: 98-377279/32.  
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
PS useful for controlling microbial infestations of plants or mammals  
PS Claim 1; Page 60-62; 96pp; English.  
CC The sequence is that of an antimicrobial protein which can  
CC be used to control microbial infestations in plants and mammalian  
CC animals.  
SQ Sequence 637 AA;  
Query Match 28.5%; Score 141; DB 1; Length 637;  
Best Local Similarity 40.7%; Pred. No. 9, 77e-05;  
Matches 22; Conservative 12; Mismatches 18; Indels 2; Gaps 2;  
Db 42 QOCVORCQERPRYSHARCVOECRDQOQHGRHEEFGRG-RGMHGEGEREE 94  
QY 195 EDCRRRCQEQEPRQHQ--QCQLRCRQEQRHGRGDMNPNRGSGRYEGEEDQ 247  
RESULT 8  
ID W62835 standard; Protein; 593 AA.  
AC W62835;  
DE 27-OCT-1998 (first entry)  
DT Zea mays antimicrobial protein.  
KW antimicrobial protein; infestation; control.  
OS Zea mays.  
PN W09827805-A1.  
PD 02-JUL-1998.  
PF 22-DEC-1997; AU0874.  
PR 20-DEC-1996; AU-004275.  
PI (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP.  
DR WPI: 98-377279/32.  
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
PS useful for controlling microbial infestations of plants or mammals  
PS Claim 1; Page 58-60; 96pp; English.  
CC The sequence is that of an antimicrobial protein which can  
CC be used to control microbial infestations in plants and mammalian  
CC animals.  
SQ Sequence 593 AA;  
Query Match 23.3%; Score 115; DB 1; Length 593;  
Best Local Similarity 31.5%; Pred. No. 2, 17e-02;  
Matches 17; Conservative 19; Mismatches 15; Indels 3; Gaps 3;  
Db 39 QCVRCEDR-PWHQRPCLCEQREERERKROERSRHEADRSGESSSEDEROE 91  
QY 196 DCRRCQEQEPRQHQ--QCQLRCRQEQRHGRGDMNPNRGSGRYEGEEDQ 247  
RESULT 9  
ID R12223 standard; Protein; 918 AA.  
AC R12223;  
DE 20-AUG-1991 (first entry)  
DT Human androgen receptor.  
KW hAR; DNA-binding protein; steroid hormone.  
OS Homo sapiens.  
PI Key Location/Qualifiers  
FT domain 556..626  
FT /label= "DNA-binding domain  
FT /note= "cysteine-rich"  
FN W09107423-A.  
PN 30-MAY-1991.  
PD 19-OCT-1990; U06015.  
PF 17-NOV-1989; US-438775.  
PR (ARCH-) ARCH DEV CORP.  
PI Liao S, Chang C.  
DR WPI: 91-178048/24.  
DR N-PSDB; Q12001.  
PT Androgen receptor and TR2 DNA binding proteins - DNA sequences  
PT and antibodies for detection and quantification methods  
PS Claim 25; Fig 3; 79pp; English.  
CC This sequence was deduced from a cDNA clone isolated by screening  
CC commercially available human testis and prostate lambda gII cDNA  
CC libraries. The sequence is very similar to that of rat AR and in  
CC the DNA-binding domain it is identical to that of rat DNA-binding  
CC domain. Homology comparisons with other known steroid receptors  
CC indicate that hAR is more closely related to glucocorticoid,  
CC mineralo-corticoid and progesterone receptors than to v-erb-A or to  
CC receptors for oestrogen, vitamin D and thyroid hormones.  
SQ Sequence 918 AA;  
Query Match 22.3%; Score 110; DB 1; Length 918;  
Best Local Similarity 39.1%; Pred. No. 5, 99e-02;  
Matches 18; Conservative 11; Mismatches 16; Indels 1; Gaps 1;  
Db 64 QQQQQQQQQQETSPRQOQQQGGGSPQARRGPTG-YTLWDEEO 108  
QY 202 EQQPRQHQCLRCRQEQRHGRGDMNPNRGSGRYEGEEDQ 247  
RESULT 10  
ID W14783 standard; Protein; 919 AA.  
AC W14783;  
DE 22-JUN-1997 (first entry)  
DT Androgen receptor.  
KW Androgen receptor; acidic fibroblast growth factor; aFGF;  
KW antisense; benign prostatic hyperplasia; prostate cancer; therapy.  
OS Homo sapiens.  
PN W09711170-A1.  
PD 27-MAR-1997.  
PF 20-SEP-1996; U15081.  
PR 20-SEP-1995; US-004018.  
PI (WOCR-) WORCESTER FOUND BIOMEDICAL RES.





